NASA Nebula Cloud Computing Platform

Cloud Computing for a Universe of Data

nebula.nasa.gov ☒ Transparency ☒ Participation ☒ Collaboration

Nebula is an open-source cloud computing platform that was developed to provide an easily quantifiable and improved alternative to building additional expensive data centers and to provide an easier way for NASA scientists and researchers to share large, complex data sets with external partners and the public. Nebula is currently an Infrastructure-as-a-Service (IaaS) implementation that provides scalable compute and storage for science data and Web-based applications. Nebula IaaS allows customers to unilaterally provision, manage, and decommission computing capabilities (virtual machine instances, storage, etc.) on an as-needed basis through a Web interface or a set of command-line tools. Platform-as-a-Service (PaaS), scheduled for release in Q4 2010, will provide a consistent set of frameworks, code repositories, and Web services to enable NASA developers to deploy secure, policy-compliant Web applications that automatically scale to meet variable demand. Software-as-a-Service (SaaS) and Database-as-a-Service (DBaaS) are planned for 2011.

Overview

Nebula emerged in 2008 out of NASA Ames Research Center. Nebula allows NASA to realize significant cost savings through better resource utilization, reduced energy consumption, and by reducing the labor required to procure infrastructure or create new Web applications. Many of NASA's dedicated compute and storage servers are underutilized but still require expensive environmental controls and a high level of on-going energy investment. Nebula allows NASA scientists to pool IT resources, only using what services they need for the time period they need it, and enabling those resources to be used by others when they no longer need them.

Nebula's high-density architecture allows for a dramatically reduced data center footprint. Each shipping container data center can hold up to 15,000 CPU cores or 15 petabytes (one petabyte equals one million gigabytes), proving 50 percent more energy efficient than traditional data centers. In addition, this "green" architecture allows for maximum flexibility and efficiency since these modular shipping container data centers can be modified, upgraded, expanded and even physically relocated as NASA's computing needs evolve over time.

Nebula Container at NASA Ames Research Center

In a traditional IT environment, it takes several months and usually hundreds of hours of labor by several different people to procure, set up, configure, and maintain new IT infrastructure. NASA must comply with a host of data security and privacy policies, which can sometimes create a challenge in finding a collaborative environment in which to share data with its outside partners. By utilizing Nebula, users gain access to powerful IT resources months faster

and with far less effort than before. Nebula saves hundreds of staff hours, allowing NASA scientists to focus on mission-critical activities instead of IT infrastructure requirements. Nebula's architecture is designed from the ground up for interoperability with commercial Cloud service providers such as Amazon Web Services, offering NASA researchers the ability to easily port data sets and code to run on commercial Clouds.

To facilitate the adoption of cloud computing within NASA and the Federal Government, Nebula team members serve on various Federal cloud working groups. Nebula's sponsor, Ames CIO Chris C. Kemp, chairs the Federal Cloud Computing Standards Working Group and the two security specialists responsible for Nebula's IT Security Plan Certification and Accreditation (CandA) are active members of the Federal Cloud Computing Security Working Group. Nebula's participation with such groups will help both NASA and the Federal Government realize the many benefits of cloud computing on an accelerated timetable.

Nicknamed the "Super Cloud," Nebula can effortlessly manage 10,000 or 100,000 times the amount of information as the most powerful commercial cloud computing platforms, accommodating files as large as eight terabytes and accommodating an individual file system of 100 terabytes (one terabyte equals 1,000 gigabytes). By contrast, the maximum Amazon EC2 file size and file system size is one terabyte. Built upon a converged 10Gig-E switching fabric, Nebula delivers 10 times the networking speed of the fastest available commercial cloud environments, most of which run at 1GigE, and use only 100Mb. This combination of high-speed networking, 2.9GHz CPUs, and hardware RAID configurations allows the Nebula environment to provide massively parallel performance equivalent to the best dedicated hardware currently available, and far in excess of any commercial cloud.

Cloud computing is a significant departure from the traditional IT infrastructure model. It will require technical training as well as a shift in how our teams and programs think about IT resources. This shift will require us to change budgeting, procurement, workflow, and our approach toward processing, storing, and accessing data. As a mission-driven Agency that is reliant upon data, we will need to provide adequate education and training to employees and contractors for greater adoption.

How This Fits into Open Government

The NASA Open Government Framework states that agencies must address issues surrounding policy, tools, and culture if they want to successfully implement Open Government principles within their organization. Nebula is the key tool that will enable the easy flow of information between NASA and the public and is the cornerstone of this Open Government Plan.

NASA Nebula hosts USAspending.gov 2.0

www.usaspending.gov

USAspending.gov 2.0 is hosted on NASA Nebula

Nebula partnered with the General Services Administration (GSA) to host USAspending.gov 2.0.

Nebula IaaS makes it easier, faster, more secure and much less expensive to deliver data-driven Web sites that encourage public participation and collaboration. Nebula allows groups inside NASA to provision resources within minutes, avoiding the lengthy procurement, certification, and security processes required for new computing infrastructure.

NASA currently has more than 3,000 external-facing Web sites deployed on variety of different platforms. Nebula PaaS will help NASA create a more coherent Web experience for the public by allowing smaller projects and offices to use the same platform as the large programs and directorates. It will also provide NASA software engineers with a robust development environment and sophisticated set of tools that make it faster, easier, and much less expensive to deliver data-driven and scalable Web sites that encourage public participation and collaboration.

NASA and Microsoft's World Wide Telescope

www.microsoft.com/presspass/press/2009/mar09/03-24NASADataPR.mspx

WorldWide Telecope's access to NASA's Mars Data stored on NASA Nebula

Through a Space Act Agreement, NASA computer scientists are working with Microsoft engineers to make NASA's data accessible through Microsoft's World Wide Telescope platform. NASA Nebula processed and hosted more than 100 terabytes of data for this project, completing what would have taken at least four to five months in just three days.

Nebula opens the doors to crowd sourcing and collaboration with powerful, economical computing resources that are built for government. The flexible capability that Nebula offers hastens the pace of innovation, collaboration, and new breakthroughs in a way that we see everyday in the private sector. By working with the open source community and operating in a fully transparent manner, Nebula continues to build upon NASA's heritage of forging new ground and sharing its results with other government Agencies.

Open Government Goals

- Three months
  - Complete plans to incorporate Nebula into NASA's data center strategy or IT framework as determined by the OCIO.
  - Continue testing Nebula IaaS alpha performance by supporting at least forty internal NASA projects.
  - Release Nebula IaaS beta.
  - Have a Nebula PaaS prototype by July 2010.
- Six months
  - Complete Nebula's IaaS 1.0

  - Release Nebula PaaS alpha by the end of 2010.
  - Release first package of open source contributions.
  - Increase Nebula compute and storage capacity.
- One year
  - Release Nebula PaaS beta.
  - Complete Nebula PaaS 1.0.
  - Complete Nebula IaaS 2.0.
  - Release second package of open source contributions.
  - Increase Nebula compute and storage capacity.
- Two years
  - Deploy the entire Nebula baseline suite of services (IaaS, PaaS, DBaaS and SaaS) into full production operation.
  - Enhance and augment Nebula services and capacity as part of normal lifecycle management.
  - Release additional open source contribution packages.

Useful Links

1. NASA Nebula: nebula.nasa.gov/about
2. Article from *Federal Computer Week*: fcw.com/articles/2009/12/10/open-government-cloud-computing.aspx?sc_lang=en
3. Article from *Government Computing News*: gcn.com/articles/2010/02/15/nasa-nebula-sidebar.aspx
4. Article from Govfresh.tv: govfresh.com/2010/02/nasa-nebula-sends-government-to-the-cloud/
5. Article from *Government Technology*: www.govtech.com/gt/717731